1971-03

Investigation of a fit of beta and normal distributions to a product of beta distribution.

Foard, John Stager Jr.

Monterey, California ; Naval Postgraduate School

http://hdl.handle.net/10945/15773

INVESTIGATION OF FIT OF BETA AND NORMAL DISTRIBUTIONS TO A PRODUCT OF BETA DISTRIBUTION

by

John Stager Foard

United States
Naval Postgraduate School

THESIS

INVESTIGATION OF FIT OF BETA AND NORMAL
DISTRIBUTIONS TO A PRODUCT OF BETA DISTRIBUTION

by

John Stager Foard, Jr.

Thesis Advisor: W. M. Woods

March 1971

Investigation of Fit of Beta and Normal Distributions to a Product of Beta Distributions

by

John Stager Foard, Jr.
Lieutenant, United States Navy
B.S., United States Naval Academy, 1965

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
March 1971

ABSTRACT

Excessive costs prohibit the Department of Defense and industry from testing entire systems to obtain estimates of system reliability. Thus, individual component tests must be combined in such a way as to yield an accurate estimate of the system reliability.

Assuming each component reliability to have a prior distribution which is beta distributed, Bayesian techniques result in the posterior distributions also being beta distributed. The system reliability for a series system would then be the product of the posterior distributions.

Computer simulation techniques were used to determine the system reliability distribution for a series system of beta distributed components. Method of Moments techniques were then used to fit beta and normal distributions to the system distribution. The twentieth percentile points of the fitted and the actual distributions were then compared as a measure of accuracy of the fit.

The fit of the beta distribution proved to be accurate for all parameter ranges and number of components in each system. The fit of the normal distribution was accurate only when used with limited parameter values.

TABLE OF CONTENTS

LIST OF TABLES

I. INTRODUCTION AND SUMMARY

A. INTRODUCTION

The purpose of this study was to analyze the accuracy of fitting a product of beta distributions with beta and with normal distributions using computer simulation techniques. If successful, a normal or beta distribution could be used to fit the posterior distributions in Bayesian techniques to the reliability of a series system where the component prior distributions are also beta distributed.

B. MEASURE OF ACCURACY

The Bayesian $100(1-\alpha)\%$ lower confidence limit for the system reliability, R_s, is the $100\ \alpha^{th}$ percentile point of the distribution of R_s. Therefore, the difference between the $100\ \alpha^{th}$ percentile points of the actual and the fitted distributions was selected as a measure of the accuracy of the fit.

C. SUMMARY OF RESULTS

When the method of moments technique was used to fit a normal distribution to the distribution of the reliability of a series system, the fit was accurate for a limited number of cases. The data depicted in Table III indicates that the normal distribution fit can offer a simple and accurate method of estimating the system reliability.

When the constraints of the fitted normal cannot be met or if an easy method of obtaining a percentile point

of a beta distribution is available, the beta fit should be used.

II. THE PROBLEM

A. BACKGROUND

Due to the costliness of mission testing entire systems, the Department of Defense must rely on individual component tests to estimate system reliability. Among other things, the accuracy of this estimate is dependent upon the fit of a specific distribution to actual distributions. It is often necessary due to mathematical and/or statistical complications to assume a distribution which only approximates the actual distribution. A secondary yet important consideration is the manner in which the component test data is utilized in determining the estimated system reliability.

Bayesian techniques offer an acceptable solution to the test data utilization. In Bayesian analysis, the beta distribution with parameters a and b is usually choosen as the prior distribution for component reliability. The parameters a and b relate to attribute type data. A problem arises when trying to combine the component beta distributions to get a confidence interval on system reliability. In particular, the problem exists on how to construct methods for estimating the system reliability when each of the component reliabilities is assumed to be beta distributed.

B. BETA DISTRIBUTION AND BAYESIAN TECHNIQUES

1. Beta Distributions

The beta distribution with parameters a and b has the following probability function:

$$f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} a,b,x>0.$$

The mean, μ, and variance, σ^2, are computed by the following formulae:

$$\mu = \frac{a}{a+b}$$

$$\sigma^2 = \frac{ab}{(a+b) \quad (a+b+1)}$$

2. Bayesian Techniques

The usefulness of beta prior distributions and Bayesian techniques can be depicted by the following example. Suppose n components of one type are mission tested and further suppose each component is beta distributed, $\beta(x;a,b)$. Let s be the number of components which do not fail. The $100(1-\alpha)\%$ lower Bayesian confidence limit on component reliability is the solution for P in the equation

$$\alpha = \int_0^P \beta(x;a+s,a+b+n-s+1)dx.$$

That is, the posterior distribution of component reliability is $\beta(x;a+s,b+n-s)$. Thus, by the use of beta distributions and Bayesian techniques, both past and present test data can readily utilized in reliability calculations.

III. SOLUTION

A. ACTUAL SYSTEM DISTRIBUTION

The actual system distribution to which the beta and normal distributions were fitted was generated on the IBM 360 Computer at the United States Naval Postgraduate School, Monterey, California. The main program was used with slight modifications to evaluate the $100\alpha^{th}$ percentile points of the fitted beta distributions. The required subroutines were obtained from the Computer Center Library.

1. Main Program

The main program was constructed to randomly generate a reliability value for each component of the series systems. Assuming the component distributions to be independent, the system reliability, R_s, would then be the product of the component reliabilities, R_i; i.e., $R_s = \prod_{i=1}^{K} R_i$. Thus, by multiplying the generated component reliability values, a value of the system reliability was obtained. This process was repeated to obtain 500 values of R_s. The values were ordered to obtain a simulated distribution of R_s. The one hundredeth number of this distribution was then assumed to be the twentieth percentile point of the actual system reliability distribution.

2. Subroutines

The Subroutine BDTR was used to obtain values of the cumulative distribution function of the beta distributions. The subroutine is contained in R. E. Bargmann and S. P. Ghosh,

Statistical Distribution Programs for a Computer Language, I.B.M. Research Report RC-1094. It has a maximum error of 10^{-5} for values of a and b between 0.5 and 10^5. Subroutine BDTR used subroutine NDTR, CDTR, and DLGAM in its computations.

B. METHOD OF MOMENTS BETA FIT

The supposition is that the system reliability, R_s, is beta distributed with parameters A and B where $R_s = \prod_{i=1}^{K} R_i$ and the R_i's are beta distributed with parameters a_i and b_i respectively. Using the method of moments technique, the mean, μ, and variance, σ^2, of the system distribution are as follows:

$$\mu = E(R_s) = \prod_{i=1}^{K} E(R_i)$$

$$= \frac{A}{A+B} = \prod_{i=1}^{K} \frac{a_i(a_i+1)}{(a_i+b_i)(a_i+b_i+1)} .$$

$$\sigma^2 = \frac{AB}{(A+B)^2(A+B+1)} = E(R_s)^2 - [E(R_s)]^2 .$$

In addition,

$$E(R_s^{\,2}) = \frac{A(A+1)}{(A+B)(A+B+1)} = \prod_{i=1}^{K} \frac{a_i(a_i+1)}{(a_i+b_i)(a_i+b_i+1)} .$$

Letting

$$\mu_1 = \prod_{i=1}^{K} \frac{a_i}{a_i+b_i}$$

and

$$\mu_2 = \prod_{i=1}^{K} \frac{a_i(a_i+1)}{(a_i+b_i)(a_i+b_i+1)} ,$$

A and B can be determined:

$$A = \frac{(\mu_1 - \mu_2)\ \mu_1}{\mu_2 - \mu_1{}^2}$$

$$B = \frac{(1-\mu_1)(\mu_1 - \mu_2)}{\mu_2 - \mu_1{}^2}$$

Subroutine BDTR was then used to determine the twentieth percentile point of the beta distribution with parameters A and B as calculated.

C. METHOD OF MOMENTS NORMAL FIT

Define $Q_i = 1 - R_i$ and assume R_i is large." Q_i is a beta distributed random variable with parameters b_i and a_i. R_s is equal to $\prod_{i=1}^{K}(1-Q_i)$. Since Q_i is "small," expanding $\prod_{i=1}^{K}(1-Q_i)$ and neglecting the cross products, $R_s = 1 - \prod_{i=1}^{K} Q_i$ and $1 - R_s = \prod_{i=1}^{K} Q_i$. The supposition is now that $1-R_s$ is normally distributed, with parameter μ and σ^2 since it is a sum of random variables. Using method of moments techniques,

$$\mu = E(1-R_s) = E(\prod_{i=1}^{K} Q_i) = \prod_{i=1}^{K} E(Q_i) = \prod_{i=1}^{K} \frac{b_i}{a_i + b_i}$$

and

$$\sigma^2 = var(1-R_s) = var(\prod_{i=1}^{K} Q_i) = \prod_{i=1}^{K} var(Q_i)$$

$$= \prod_{i=1} \frac{a_i b_i}{(a_i + b_i)^2\ (a_i + b_i + 1)} .$$

The twentieth percentile point of R_s is the same as the eightieth percentile point of $1-R_s$. The eightieth percentile point of $1-R_s$ is $\sigma Z_{.8} + \mu$ where $Z_{.8}$ is the eightieth percentile point of the Standard Normal Distribution.

IV. RESULTS

A. ACTUAL SYSTEM DISTRIBUTION

The actual system distributions were generated for series systems of fifteen and twenty-five components using the nineteen sets of parameters as listed in the first three tables. It was assumed that all components in the systems would have the same parameter values. The twentieth percentile points of these distributions are listed in Tables I and II.

Due to the number of runs required and to the length of computer time required to call subroutine BDTR, the number of samples forming these distributions was limited to five hundred. This limitation resulted in the possible variance of the twentieth percentile point of a given distribution by as much as .003 when a different series of random numbers were used to generate the same distribution. Variations of this amount would make a significant change in column five of Tables I and II. For example, using a=1000.0 and b=5.0 in Table I, the difference between the actual twentieth percentile point and the fitted beta point was only .002. Yet, this difference resulted in an error of 4.6 percentage points (20.0-15.4=4.6).

B. METHOD OF MOMENTS BETA FIT

Using the formulae developed previously to obtain the system parameters A and B, the twentieth percentile point of the system distribution was calculated by computer using

subroutine BDTR. Considering the maximum error of BDTR, the data of column four in Tables I and II should be more accurate than that of the actual distribution.

Examination of columns three and four of Tables I and II shows that the beta fitted distribution provides a more accurate estimate of the actual distribution for the system of twenty-five components than for the fifteen. However, even for the smaller system, the method does result in reasonably accurate estimates of the twentieth percentile point particularly if a variance of .003 is introduced into the actual distribution percentile point. It is interesting to note that the method appears to be about the same degree accurate throughout the parameter ranges for both systems. Due to the steepness of the curves involved, the small percentage error figures can be misleading. The actual percentile points (column five) might offer a better measure of accuracy. These points indicate that in most of the cases the error is conservative and that the approximation actually represents the lower limit of a higher degree of confidence than eighty percent.

C. METHOD OF MOMENTS NORMAL FIT

The results of the normal fit are depicted in Table III. As can readily be seen in this table, the normal fit is not accurate when the ratio of a to b is less than 500 to 1. Since these parameters can be related to the number of failures and number of successes, these results tend to prove the initial assumption made when deriving this method.

This assumption was that R_i was "large." The results indicate that "large" means one or less failures out of 500 mission tested components. Given these parameter constraints, the normal fit is quite accurate. Due to the limited range of parameters, this method was not investigated further.

D. BETA FIT WITH VARIOUS PARAMETERS

The process of fitting the beta distribution to the system reliability was successful enough to investigate the accuracy of the method when the components have different parameter values. Table IV depicts these results. The comparison was made for systems of fifteen and twenty-five components and for two different general parameter ranges. The limitations of computer core suggested that no more than five different distributions be used. Thus, in the fifteen component system, there are three components with each set of parameters; and in the twenty-five component system, there are five components with each set of parameters.

The results indicate that the method of moment fitted beta distribution is equally accurate when there are mixed distributions in the systems.

TABLE I

This table lists the results of the method of moments beta fit and the actual distribution for a series system of fifteen components. All distributions in the system have the same parameters. Columns one and two list the a and b parameters respectively. Column three is the twentieth percentile point of the actual distribution. Column four is the twentieth percentile point of the beta fitted distribution. Column five lists the percentile point of the actual distribution that the number in column four represents. Column six is the percent error which is the difference between column three and column four divided by column three.

a	b	Actual	Beta Fit	Act.%	Error %
10.0	0.5	0.372	0.372	20.0	0.0
10.0	1.0	0.168	0.163	18.8	3.0
30.0	0.5	0.768	0.723	4.4	5.9
30.0	1.0	0.545	0.547	20.8	0.3
30.0	2.0	0.328	0.323	17.4	1.6
50.0	0.5	0.825	0.824	19.4	0.1
50.0	1.0	0.700	0.696	18.8	0.6
50.0	2.0	0.529	0.505	7.8	4.5
100.0	0.5	0.913	0.908	16.4	0.6
100.0	1.0	0.839	0.834	17.6	0.6
100.0	2.0	0.716	0.709	16.4	1.0
500.0	0.5	0.981	0.981	20.1	0.0
500.0	1.0	0.965	0.964	17.8	0.0
500.0	2.0	0.936	0.933	13.0	0.5
500.0	5.0	0.850	0.849	17.4	0.2
1000.0	0.5	0.990	0.990	20.1	0.0
1000.0	1.0	0.983	0.982	17.6	0.0
1000.0	2.0	0.967	0.966	17.6	0.0
1000.0	5.0	0.923	0.921	15.4	0.2

TABLE II

This table lists the results of the method of moments beta fit and the actual distribution for a series system of twenty-five components. All distributions in the system have the same parameters. Columns one and two list the a and b parameters respectively. Column three is the twentieth percentile point of the actual distribution. Column four is the twentieth percentile point of the beta fitted distribution. Column five lists the percentile point of the actual distribution that the number in column four represents. Column six is the percent error which is the difference between columns three and four divided by column three.

a	b	Actual	Beta Fit	Act. %	Error %
10.0	0.5	0.205	0.207	20.8	0.8
10.0	1.0	0.055	0.054	19.4	1.8
30.0	0.5	0.650	0.597	2.6	8.2
30.0	1.0	0.379	0.379	20.0	0.0
30.0	2.0	0.160	0.160	20.1	0.0
50.0	0.5	0.733	0.735	20.8	0.3
50.0	1.0	0.560	0.559	19.6	0.2
50.0	2.0	0.358	0.331	5.0	7.5
100.0	0.5	0.860	0.857	17.6	0.3
100.0	1.0	0.747	0.747	20.0	0.0
100.0	2.0	0.575	0.574	19.4	0.2
500.0	0.5	0.970	0.970	20.8	0.0
500.0	1.0	0.944	0.944	19.6	0.0
500.0	2.0	0.897	0.894	14.2	0.3
500.0	5.0	0.765	0.765	21.0	0.0
1000.0	0.5	0.985	0.985	21.0	0.0
1000.0	1.0	0.971	0.971	19.0	0.0
1000.0	2.0	0.945	0.946	20.8	0.1
1000.0	5.0	0.875	0.875	18.0	0.0

TABLE III

This table lists the results of the method of moments normal fit and the actual distribution for a series system of twenty-five components. All distributions in the system have the same parameters. Columns one and two list the a and b parameters respectively. Column three is the twentieth percentile point of the actual distribution. Column four is the twentieth percentile point using the techniques described previously.

a	b	Actual	Normal Fit
10.0	0.5	0.205	- 0.454
10.0	1.0	0.055	- 1.62
30.0	0.5	0.650	0.49
30.0	1.0	0.379	0.062
30.0	2.0	0.160	- 0.739
50.0	0.5	0.733	0.694
50.0	1.0	0.560	0.429
50.0	2.0	0.358	- 0.072
100.0	0.5	0.860	0.846
100.0	1.0	0.747	0.711
100.0	2.0	0.575	0.452
500.0	0.5	0.970	0.969
500.0	1.0	0.944	0.942
500.0	2.0	0.897	0.888
500.0	5.0	0.765	0.734
1000.0	0.5	0.985	0.984
1000.0	1.0	0.971	0.971
1000.0	2.0	0.945	0.944
1000.0	5.0	0.875	0.866

TABLE IV

This table lists the results for method of moments beta fit using different parameters in the component distributions. Columns one and two are a and b parameters used. Column three is the number of components in the system. Five different sets of parameters are used. In the fifteen component system, there are three components with the same set of parameters. In the twenty-five component system, there are five components with the same set of parameters. Column four is the twentieth percentile point of the actual distribution. Column five is the twentieth percentile point of the fitted beta distribution.

a	b	k	Actual	Beta Fit
50.0	0.5			
100.0	0.5	15	0.822	0.820
100.0	1.0			
100.0	2.0	25	0.739	0.728
500.0	5.0			
500.0	1.0			
500.0	2.0	15	0.966	0.966
1000.0	0.5			
1000.0	1.0	25	0.946	0.947
1000.0	2.0			

TABLE V

Using the method of moments beta fit, R_s was assumed to be beta distributed with parameters A and B. This table lists the parameters A and B for the various numbers of system components and component parameters a and b. Columns one and two are the a and b parameters respectively. Columns three and six are the number of components in the system. Columns four and five are the A and B values for the variables of columns one, two, and three. Columns seven and eight are the A and B values for the variables of columns one, two, and six.

1.a	2.b	3.k	4.A	5.B	6.k	7.A	8.B
10.0	0.5	15	7.2	7.5	25	5.9	14.0
10.0	1.0	15	5.5	17.5	25	3.8	37.8
30.0	0.5	15	26.8	7.5	25	24.8	12.7
30.0	1.0	15	24.1	15.3	25	20.8	26.4
30.0	2.0	15	19.8	32.3	25	15.3	61.4
50.0	0.5	15	46.7	7.5	25	44.5	12.6
50.0	1.0	15	43.7	15.1	25	39.8	25.5
50.0	2.0	15	38.5	30.9	25	32.4	54.0
100.0	0.5	15	96.6	7.5	25	94.3	12.5
100.0	1.0	15	93.3	15.0	25	88.9	25.1
100.0	2.0	15	87.4	30.2	25	79.6	51.0
500.0	0.5	15	500.7	7.6	25	498.9	12.6
500.0	1.0	15	490.7	14.9	25	485.4	24.9
500.0	2.0	15	485.9	30.0	25	476.1	50.0
500.0	5.0	15	465.3	74.9	25	443.4	125.2
1000.0	0.5	15	1029.9	7.8	25	1046.6	13.2
1000.0	1.0	15	976.0	14.7	25	969.2	24.5
1000.0	2.0	15	993.8	30.2	25	981.4	50.3
1000.0	5.0	15	966.8	75.1	25	942.9	125.2
50.0	0.5	3			5		
100.0	0.5	3			5		
100.0	1.0	3	89.98	16.03	5	85.27	26.79
100.0	2.0	3			5		
500.0	5.0	3			5		
500.0	1.0	3			5		
500.0	2.0	3			5		
1000.0	0.5	3	578.16	16.69	5	576.65	28.02
1000.0	1.0	3			5		
1000.0	2.0	3			5		

MAIN PROGRAM FOR FIVE DIFFERENT COMPONENT TYPES

```
      DIMENSION RSHAT(500),XARAY(5,750),PARAY(5,750),NUM(5)
      Z=URN(0)
      DATA NITER,I37,I3,KCOMP,PROD,NREP,ALPER,I31/
      DO 555 I32=1,5
  400 READ(5,401)ALFA,BETA,NSTAR,NSTOP,N,CFACT
  401 FORMAT(2F6.1,2I7,I1,F8.7)
      I2=1
      DO 402 I21=NSTAR,NSTOP,N
      X=I21*CFACT
      CALL BDTR(X,ALFA,BETA,P,DOUT,IER)
      XARAY(I32,I2)=X
      PARAY(I32,I2)=P
      I2=I2+1
  402 CONTINUE
      NUM(I32)=I2-1
  555 CONTINUE
      DO 222 I1=1,NITER
      I32=1
      DO 333 I4=1,KCOMP
      FLY=URN(1)
   49 IF(FYL-PARAY(I32,I3))50,50,51
   51 I3=I3+1
      IF(I3.EQ.NUM(I32))GO TO 60
      GO TO 49
   50 A=PARAY(I32,I3)-FYL
      IF(I3.EQ.1)GO TO 60
      IT3=I3-1
      B=FYL-PARAY(I32,IT3)
      IF(A-B)60,60,61
   61 RIHAT=XARAY(I32,IT3)
      GO TO 62
   60 RIHAT=XARAY(I32,I3)
   62 PROD=PROD*RIHAT
      I3=1
      AI4=I4
      AM=AI4/I37
      MM=AM+.2
      M=I4/I37
      IF(MM.EQ.M)I32=I32+1
  333 CONTINUE
      RSHAT(I1)=PROD
      PROD=1.
  222 CONTINUE
      J=0
    8 IF(J.EQ.NITER-1)GO TO 11
      I=1
      J=0
    9 IF (RSHAT(I).GT.RSHAT(I+1))GO TO 10
      J=J+1
      I=I+1
      IF(I.EQ.NITER)GO TO 8
      GO TO 9
   10 RTEMP=RSHAT(I)
      RSHAT(I)=RSHAT(I+1)
      RSHAT(I+1)=RTEMP
      I=I+1
      IF(I.EQ.NITER)GO TO 8
      GO TO 9
   11 CONTINUE
      STOP
      END
```

BIBLIOGRAPHY

1. Larson, H. J., Introduction to Probability Theory and Statistical Inference, Wiley, 1969.

2. Lloyd, D. K., and Lipow, M., Reliability: Management, Methods, and Mathematics, Prentice Hall, 1962.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Documentation Center Cameron Station Alexandria, Virginia 22314	2
2.	Library, Code 0212 Naval Postgraduate School Monterey, California 93940	2
3.	Assoc Professor W. M. Woods, Code 55 [Wo] Department of Operations Analysis Naval Postgraduate School Monterey, California 93940	1
4.	Department of Operations Analysis Code 55 Naval Postgraduate School Monterey, California 93940	1
5.	LT John S. Foard, Jr., USN Box 9004 Pensacola Beach, Florida 32561	1

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Naval Postgraduate School Monterey, California 93940	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP

3. REPORT TITLE

Investigation of Fit of Beta and Normal Distributions to a Product of Beta Distributions

4. DESCRIPTIVE NOTES (Type of report and, inclusive dates)

Master's Thesis, March 1971

5. AUTHOR(S) (First name, middle initial, last name)

John S. Foard, Jr.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
March 1971	24	2

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO.	
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	

10. DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Naval Postgraduate School Monterey, California 93940

13. ABSTRACT

Excessive costs prohibit the Department of Defense and industry from testing entire systems to obtain estimates of system reliability. Thus individual component tests must be combined in such a way as to yield an accurate estimate of the system reliability.

Assuming each component reliability to have a prior distribution which is beta distributed, Bayesian techniques result in the posterior distributions also being distributed. The system reliability for a series system would then be the product of the posterior distributions.

Computer simulation techniques were used to determine the system reliability distribution for a series system of beta distributed components. Method of Moments techniques were then used to fit beta and normal distributions to the system distribution. The twentieth percentile points of the fitted and the actual distributions were then compared as a measure of accuracy of the fit.

The fit of the beta distribution proved to be accurate. for all parameter ranges and number of components in each system. The fit of the normal distribution was accurate only when used with limited parameter values.

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Beta Distributions						
Bayesian Techniques						
Method of Moments Techniques						

126436
Thesis
F528 Foard
c.1 Investigation of fit of beta and normal distributions to a product of beta distribution.